Vol. 197, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 30, 1993 Pages 187-191

EFFECT OF BASE-MISMATCH IN THE COHESIVE ENDS OF OLIGONUCLEOTIDE IN GENE CLONING

Wing L. Sung*, Diana M. Zahab, Catherine K. Luk
and Benedict S. Chan

Institute for Biological Sciences, National Research Council of Canada, Ottawa, ON, Canada K1A 0R6

Received October 5, 1993

SUMMARY: The tolerance of mismatched nucleotides between the cohesive ends of insert and target DNAs in gene cloning has been investigated. An oligonucleotide duplex with a cohesive end GGCC-5′ or variation was ligated to the 5′-CCGG end of a linearized plasmid. The ligation mixture was used in the transformation of *E. coli*. A single-base mismatch, such as $^{5'\text{-CCGG}}/_{\underline{A}\text{GCC-}}$, $_{\text{G}\underline{A}\text{CC-}}$, $_{\text{GG}\underline{A}\text{C-}}$ or $_{\text{GGC}\underline{A}\text{-5'}}$ (mismatch underlined), was well tolerated in the cloning of the oligonucleotide duplex, with efficiency lower than the fully complementary ends. Double-base mismatch $^{5'\text{-CCGG}}/_{\underline{AA}\text{CC-}}$ or $_{\text{GG}\underline{AA}\text{-5'}}$ resulted in further decrease of cloning efficiency. Via a similar approach, a tetracycline resistance gene was successfully inserted into a pUC-type plasmid.

DNA ligation is a routine process in gene cloning (1, 2). With the exception of the blunt-ends, ligation takes place at the complementary, single-stranded ends of two DNA duplexes (1, 2). Two short "cohesive ends", such as those of 4 or fewer bases created by restriction endonucleases, anneal (at least momentarily) for the ligase to form covalent linkages. The stringent requirement of complementary ends has been applied in the development of an *in vitro* "ligase chain reaction" assay for the rapid detection of genetic diseases.

However, such a condition is not strictly observed in the *in vivo* process of gene cloning. Partially complementary oligonucleotides have been successfully cloned into plasmid (3). In these cases, the mismatch was some distance from the site of ligation and buried within the overlapping DNA duplex (3). One of

* To whom correspondence should be addressed.

Abbreviations: AmpR, ampicillin resistance gene; PTH, human parathyroid hormone; TetR, tetracycline resistance gene.

0006-291X/93 $4.00

our objectives in this investigation was the insertion of a tetracycline resistance TetR gene cartridge into a plasmid, so that the human parathyroid hormone PTH could be expressed by *E. coli* in a culture medium containing tetracycline, instead of ampicillin. However, such an insertion would involve ligation between the *Xma*I (5'-CCGG) and *Ava*I (AGCC-5') cohesive ends. The single-base mismatch (underlined) may potentially prevent the proper insertion of the TetR gene into the plasmid. Therefore a model study was initiated on the tolerance of sequence mismatch in the cohesive ends during gene cloning. We have demonstrated that the simple mismatch in some cohesive ends does not necessarily deter the cloning of an oligonucleotide duplex. The TetR gene was subsequently inserted into the plasmid via such a process.

MATERIAL AND METHODS

Oligonucleotides were synthesized using a 380B DNA synthesizer (Applied Biosystem). Construction of the plasmid pPTH-AA-Eco has been reported previously (4). The protocol for the construction of all the plasmids described was identical to the published procedure (5). *E. coli* strain Y1091 (Clontech, Palo Alto, CA) was used in all expressions. A tetracycline resistance gene block was purchased from Pharmacia (Dorval, Quebec). Restriction enzymes, T4 DNA ligase and polynucleotide kinase were purchased from New England Biolab (Mississauga, ON).

RESULTS AND DISCUSSION

A synthetic oligonucleotide duplex KN-1/KN-2 (Figure 1), with *Kas*I and *Nde*I cohesive ends at its termini, was ligated to the plasmid pPTH-AA-Eco which was previously linearized by the same endonucleases (Figure 2A). The resultant plasmid pPTH-KN with the newly added *Eco*RI and *Xma*I sites (Figure 2B), was used for the present model study.

Plasmid pPTH-KN was linearized with *Eco*RI and *Xma*I nucleases. The latter endonuclease has generated a 5'-CCGG cohesive end in the linearized plasmid for our model study.

KN-1

5'-GCG CCA GAG GAA TTC TTT GTG CCC GGG CA
GT CTC CTT AAG AAA CAC GGG CCC GTA T-5'

*Kas*I *Eco*RI *Xma*I *Nde*I

KN-2

Figure 1. Oligonucleotide duplex KN-1/KN-2 for the preparation of plasmid pPTH-KN.

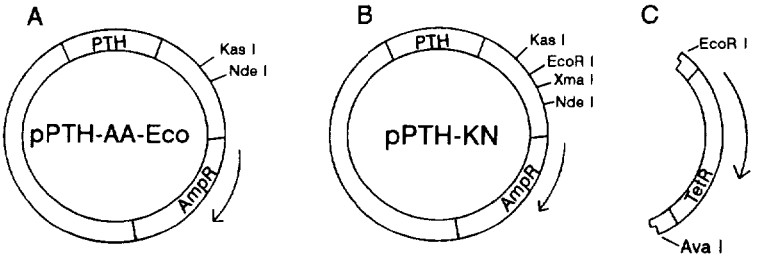

Figure 2. Plasmids pPTH-AA-Eco (A), pPTH-KN (B), and the tetracycline resistance gene insert TetR (C). Plasmid pPTH-KN was derived from pPTH-AA-Eco. The *Eco*RI/*Xma*I-linearized plasmid pPTH-KN was used for the insertion of the various sequences including the TetR insert.

Complementary oligonucleotides M-1/M-2, with the *Eco*RI and *Xma*I-type (5′-CCGG) cohesive ends at their respective 5′-ends (Table 1), were synthesized. The duplex was then inserted into the

Table 1. Cloning efficiency of oligonucleotide duplexes with variations of cohesive ends

Oligonucleotides	cloning efficiency %[c]	Binding energy of cohesive ends[d] ΔG, kcal/mol
*Eco*RI		
M-1 5′-AATTCCCCGGACGTCTT		
M-2 GGGGCCTGCAGAAGGCC-5′		
*Xma*I-type		
M-2 and variants[a]		
M-2 GGGGCCTGCAGAAGGCC-5′	62	-9.8
M-2a .....AGCC-5′	17	-6.7
M-2b .....GACC-5′	17	-3.1
M-2c .....GGAC-5′	40	-3.1
M-2d .....GGCA-5′	57	-6.7
M-2e .....AACC-5′	1	-3.1
M-2f .....GGAA-5′	8	-3.1
M-1b and variants[b]		
M-1 5′-AATTCCCCGGACGTCTT	62	-5.3
M-1a 5′-AATG.....	0	-3.4
M-1b 5′-AATC.....	0	-3.4

[a] Oligonucleotide M-2 or variant was annealed to M-1 for subsequent insertion into *Eco*RI/*Xma*I-cut plasmid pPTH-KN. Base-mismatch between the *Xma*I end of the plasmid and the 5′-end of M-2a to M-2f is underlined.

[b] Oligonucleotide M-1 or variant was annealed to M-2 for subsequent insertion into *Eco*RI/*Xma*I-cut plasmid pPTH-KN. Base-mismatch between the *Eco*RI end of the plasmid and the 5′-end of M-1a to M-1b is underlined.

[c] Cloning efficiency was calculated as the percentage of randomly selected transformants that were positively hybridized to the ^{32}P-labelled probe M-1.

[d] Binding energy ΔG between the cohesive ends of the oligonucleotide and the linearized plasmid (6).

*Eco*RI/*Xma*I-cut plasmid pPTH-KN via ligation with T4 DNA ligase (Table 1). After transformation of *E. coli*, transformants were randomly selected for hybridization with ^{32}P-labelled probe M-1. Sixty-two percent of the transformant candidates possessed the insert (Table 1).

Variants of the oligonucleotide M-2 were also prepared as substitutes for M-2 in the same insertion (Table 1). Single-base mismatches were created in the four locations in the 4-base cohesive end in M-2a, -2b, -2c, and -2d. In all cases, the new duplexes were successfully inserted into the plasmid with efficiencies ranging 17-57%.

The effect of double-base mismatch was also investigated. Duplexes M-1/M-2e and M-1/M-2f designed for such mismatch have been successfully inserted, but with lower efficiencies of 1 and 8% respectively (Table 1). Because of the very low tolerance for double-base mismatch, no attempt was made to study the effect of a triple-base mismatch.

In each experiment involving single- and double-base mismatch, plasmids of 4 positive transformants were prepared for nucleotide sequencing. The proper insertion of a single copy of the duplex was confirmed in all cases. As expected in the replication of heteroduplex plasmid, two versions of the daughter plasmid derived from either M-1 or M-2 variants have been isolated.

Tolerance of base mismatch in the A,T-rich *Eco*RI cohesive end was also investigated. In contrast to the *Xma*I cohesive end of C and G nucleotides, the *Eco*RI cohesive end in opposite terminus of the linearized plasmid is made of A and T nucleotides. Variants of M-1 with a single-base mismatch were prepared (Table 1). However, duplexes of M-1a/M-2 and M-1b/M-2 failed to be inserted into the linearized plasmid.

To demonstrate the relevance of the above findings in gene cloning, an attempt was made to insert a TetR gene between the same *Eco*RI and *Xma*I sites of the plasmid pPTH-KN. After cutting by these two endonucleases, the plasmid was ligated to a TetR gene cartridge which possessed an *Eco*RI and a *Ava*I cohesive ends (Figure 2C). Annealing between the *Xma*I (5′-CCGG) and the partially complementary *Ava*I (AGCC-5′) ends was expected to create a single base mismatch identical to that of M-1/M-2a (Table 1). After transformation in *E. coli* with ampicillin for selection, hybridization with a labelled TetR probe indicated that 70% of the transformants possessed the TetR gene, as predicted by the above model study (Table 1). Proper insertion of the TetR gene was

subsequently confirmed by nucleotide sequencing. As expected, two versions of daughter plasmid have been identified, which possess either the *AvaI* or *XmaI* sequence. Tetracycline resistance has been expressed by both plasmids.

The present study of gene insertion indicated that the C,G-rich cohesive end has tolerated base-mismatch, while the A,T-rich cohesive end failed. In the C,G-rich cohesive end, a single-base mismatch was more tolerated in contrast to the double-base mismatch. This may be the result of different binding energies ΔG in the formation of short heteroduplex between cohesive ends for the initial ligation (6). However, binding energy alone cannot sufficiently account for the efficiencies of different variants M-2 (e.g. M-2c/M-1 in Table 1). The circularization of these plasmids is unlikely to be a single-step process of *in vitro* ligation between partially complementary cohesive ends. It may involve the complicated DNA repair mechanism (7-9).

REFERENCES

1. Cohen, S. N., Chang, A. C. Y., Boyer, H. W., Helling, R. B. (1973) Proc. Natl. Acad. Sci. U.S.A. 70, 3240-3244.

2. Mertz, J. E. and Davis, R. W. (1972) Proc. Natl. Acad. Sci. U.S.A. 69, 3370-3374.

3. Sung, W. L., Zahab, D. M., Yao, F.-L., Wu, R., and Narang, S. A. (1986) Nucl. Acids Res. 14, 6159-6168.

4. Sung, W. L., Zahab, D. M., Barbier, J. R., Watson, D., Yaguchi, M., Neugebauer, W., and Willick, G. E. (1991) J. Biol. Chem. 266, 2831-2835.

5. Sung, W. L., Yao, F.-L., Zahab, D. M. and Narang, S. A. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 561-565.

6. Breslauer, K. J., Frank, R., Blocker, H., and Marky, L. A. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 3746-3750.

7. Wagner, R. and M. Meselson (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 4135-4139.

8. Lu, A.-L., Clark, S., and Modrich , P. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 4639-4643.

9. Hanawalt, P. C., Friedberg, E. C., and Fox, C. F. (eds.)(1978) DNA Repair Mechanisms, Academic Press, New York.